In re patent application of KNAPPIK, Achim et al.

Serial No.: 09/490,064

Filed: January 24, 2000

For: PROTEIN/(POLY)PEPTIDE LIBRARIES

Group Art Unit: 1631

Examiner: Mary K. Zeman

RESPONSE

COPY

Assistant Commissioner for Patents Washington, D.C. 20231

Sir:

In response to the Office Action mailed March 18, 2002, applicants submit herewith substitute pages 18, 18a, 19, 19a, 20, 20a, 21, 21a, 22, 22a, 29, 29a, 31, 31a, 32, 32a, 39, 39a, 41 and 41a. The substitute pages fully comply with 37 CFR § 1.121 and, accordingly, entry of the amendments is requested.

CONCLUSION

In view of the above remarks and amendments, it is respectfully submitted that this application is in condition for allowance. Early notice to that effect is earnestly solicited. The Examiner is invited to telephone the undersigned at the number listed below if the Examiner believes such would be helpful in advancing the application to issue.

It is believed that no fees are due with this submission of this Response. In the event any fees are required for the filing of this paper, applicants authorize the Commissioner to charge such fees to Deposit Account 08-1641.

Respectfully submitted,

April 18, 2002

Date

Paul M. Booth Reg. No. 40,244

Heller Ehrman White & McAuliffe LLP

101 Orchard Ridge Drive, Suite 300

Gaithersburg, Maryland 20878

Telephone:

(301) 721-6100

Facsimile:

(301) 721-6299

LEGENDS TO FIGURES AND TABLES

- FIG. 1: Flow chart outlining the process of construction of a synthetic human antibody library based on consensus sequences.
- FIG. 2: Alignment of consensus sequences designed for each subgroup (amino acid residues are shown with their standard one-letter abbreviation). (A) (SEQ ID NOS 28-31, respectively) kappa sequences, (B) (SEQ ID NOS 32-34, respectively) lambda sequences and (C) (SEQ ID NOS 35-41, respectively), heavy chain sequences. The positions are numbered according to Kabat (1991). In order to maximize homology in the alignment, gaps (-) have been introduced in the sequence at certain positions.
- FIG. 3: Gene sequences (SEQ ID NOS 42, 44, 46 and 48, respectively) of the synthetic V kappa consensus genes. The corresponding amino acid sequences (SEQ ID NOS 43, 45, 47 and 49 respectively) (see FIG. 2) as well as the unique cleavage sites are also shown.
- FIG. 4: Gene sequences (SEQ ID NOS 50, 52 and 54 respectively) of the synthetic V lambda consensus genes. The corresponding amino acid sequences (SEQ ID NOS 51, 53 and 55 respectively) (see FIG. 2) as well as the unique cleavage sites are also shown.
- FIG. 5: Gene sequences (SEQ ID NOS 56, 58, 60, 62, 64, 66 and 68 respectively) of the synthetic V heavy chain consensus genes. The corresponding amino acid sequences (SEQ ID NOS 57, 59, 61, 63, 65, 67 and 69 respectively) (see FIG. 2) as well as the unique cleavage sites are also shown.
- FIG. 6: Oligonucleotides (SEQ ID NOS 70-164, respectively) used for construction of the consensus genes. The oligos are named according to the corresponding consensus gene, e.g. the gene Vκ1 was constructed using the six oligonucleotides O1K1 to O1K6. The oligonucleotides used for synthesizing the genes encoding the constant domains Cκ (OCLK1 to 8) and CH1 (OCH1 to 8) are also shown.
- FIGS. 7A/B: Sequences of the synthetic genes (SEQ ID NOS 165 and 167 respectively) encoding the constant domains Cκ (A) and CH1 (B). The corresponding amino acid sequences (SEQ ID NOS 166 and 168

- respectively) as well as unique cleavage sites introduced in these genes are also shown.
- FIG. 7C: Functional map and sequence (SEQ ID NOS 169-170 respectively) of module M24 comprising the synthetic Cλ gene segment (huCL lambda).
- FIG. 7D: Oligonucleotides (SEQ ID NOS 171-176 respectively) used for synthesis of module M24.
- FIG. 8: Sequence (SEQ ID NOS 177-178 respectively) and restriction map of the synthetic gene encoding the consensus single-chain fragment VH3-Vκ2. The signal sequence (amino acids 1 to 21) was derived from the E. coli phoA gene (Skerra &

Pluckthun, 1988). Between the phoA signal sequence and the VH3 domain, a short sequence stretch encoding 4 amino acid residues (amino acid 22 to 25) has been inserted in order to allow detection of the single-chain fragment in Western blot or ELISA using the monoclonal antibody M1 (Knappik & Pluckthun, 1994). The last 6 basepairs of the sequence were introduced for cloning purposes (EcoRI site).

- FIG. 9: Plasmid map of the vector pIG10.3 used for phage display of the H3κ2 scFv fragment. The vector is derived from pIG10 and contains the gene for the lac operon repressor, lacl, the artificial operon encoding the H3κ2-gene3ss fusion under control of the lac promoter, the lpp terminator of transcription, the single-strand replication origin of the E. coli phage f1 (F1_ORI), a gene encoding β-lactamase (bla) and the ColEI derived origin of replication.
- FIG. 10: Sequencing results of independent clones from the initial library, translated into the corresponding amino acid sequences. (A) (SEQ ID NO: 179) Amino acid sequence of the VH3 consensus heavy chain CDR3 (position 93 to 102, Kabat numbering). (B) (SEQ ID NOS 180-191 respectively) Amino acid sequences of 12 clones of the 10-mer library. (C) (SEQ ID NOS 192-202 respectively) Amino acid sequences of 11 clones of the 15-mer library, *: single base deletion.
- FIG. 11: Expression test of individual library members. (A) Expression of 9 independent clones of the 10-mer library. (B) Expression of 9 independent clones of the 15-mer library. The lane designated with M contains the size marker. Both the gp3-scFv fusion and the scFv monomer are indicated.
- FIG. 12: Enrichment of specific phage antibodies during the panning against FITC-BSA. The initial as well as the subsequent fluorescein-specific sub-libraries were panned against the blocking buffer and the ratio of the phage eluted from the FITC-BSA coated well vs. that from the powder milk coated well from each panning round is presented as the "specificity factor".
- FIG. 13: Phage ELISA of 24 independent clones after the third round of panning tested for binding on FITC-BSA.

- FIG. 14: Competition ELISA of selected FITC-BSA binding clones. The ELISA signals (OD.sub.405 nm) of scFv binding without inhibition are taken as 100%.
- FIG. 15: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against FITC-BSA, translated into the corresponding amino acid sequences (SEQ ID NOS 203-218 respectively) (position 93 to 102, Kabat numbering).

- FIG. 16: Coomassie-Blue stained SDS-PAGE of the purified anti-fluorescein scFv fragments: M: molecular weight marker, A: total soluble cell extract after induction, B: fraction of the flow-through, C, D and E: purified scFv fragments 1HA-3E4, 1HA-3E5 and 1HA-3E10, respectively.
- FIG. 17: Enrichment of specific phage antibodies during the panning against β-estradiol-BSA, testosterone-BSA, BSA, ESL-1, interleukin-2, lymphotoxin-β, and LeY-BSA after three rounds of panning.
- FIG. 18: ELISA of selected ESL-1 and .beta.-estradiol binding clones.
- FIG. 19: Selectivity and cross-reactivity of HuCAL antibodies: in the diagonal specific binding of HuCAL antibodies can be seen, off-diagonal signals show non-specific cross-reactivity.
- FIG. 20: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against β-estradiol-BSA, translated into the corresponding amino acid sequences (SEQ ID NOS 219-230 respectively) (position 93 to 102, Kabat numbering). One clone is derived from the 10mer library.
- FIG. 21: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against testosterone-BSA, translated into the corresponding amino acid sequences (SEQ ID NOS 231-236 respectively) (position 93 to 102, Kabat numbering).
- FIG. 22: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against lymphotoxin-.beta., translated into the corresponding amino acid sequences (SEQ ID NOS 237-244 respectively) (position 93 to 102, Kabat numbering). One clone comprises a 14mer CDR, presumably introduced by incomplete coupling of the trinucleotide mixture during oligonucleotide synthesis.
- FIG. 23: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against ESL-1, translated into the corresponding amino acid sequences (SEQ ID NOS 245-256 respectively) (position 93 to 102, Kabat numbering). Two clones are derived from the 10mer library. One clone comprises a 16mer CDR, presumably introduced by chain elongation during oligonucleotide synthesis using trinucleotides.

- FIG. 24: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against BSA, translated into the corresponding amino acid sequences (SEQ ID NOS 257-262 respectively) (position 93 to 102, Kabat numbering).
- FIG. 25: Schematic representation of the modular pCAL vector system.
- FIG. 25a: List of restriction sites already used in or suitable for the modular HuCAL genes and pCAL vector system.
- FIG. 26: List of the modular vector elements for the pCAL vector series: shown are only those restriction sites which are part of the modular system.

- FIG. 27: Functional map and sequence (SEQ ID NO: 263) of the multi-cloning site module (MCS)
- FIG. 28: Functional map and sequence (SEQ ID NOS 264-265 respectively) of the pMCS cloning vector series.
- FIG. 29: Functional map and sequence (SEQ ID NO: 266) of the pCAL module M1 (see FIG. 26).
- FIG. 30: Functional map and sequence (SEQ ID NOS 267-268 respectively) of the pCAL module M7-III (see FIG. 26).
- FIG. 31: Functional map and sequence (SEQ ID NO: 269) of the pCAL module M9-II (see FIG. 26).
- FIG. 32: Functional map and sequence (SEQ ID NO: 270) of the pCAL module M11-II (see FIG. 26).
- FIG. 33: Functional map and sequence (SEQ ID NO: 271) of the pCAL module M14-Ext2 (see FIG. 26).
- FIG. 34: Functional map and sequence (SEQ ID NOS 272-273 respectively) of the pCAL module M17 (see FIG. 26).
- FIG. 35: Functional map and sequence (SEQ ID NOS 274-276 respectively) of the modular vector pCAL4.
- FIG. 35a: Functional maps and sequences (SEQ ID NOS 277-300 respectively) of additional pCAL modules (M2, M3, M7I, M7II, M8, M10II, M11II, M12, M13, M19, M20, M21, M41) and of low-copy number plasmid vectors (pCALO1 to pCALO3).
- FIG. 35b: List of oligonucleotides and primers (SEQ ID NOS 301-360 respectively) used for synthesis of pCAL vector modules.
- FIG. 36: Functional map and sequence (SEQ ID NOS 361-362 respectively) of the β-lactamase cassette for replacement of CDRs for CDR library cloning.
- FIG. 37: Oligo and primer (SEQ ID NOS 363-367 respectively) design for Vκ CDR3 libraries.
- FIG. 38: Oligo and primer (SEQ ID NOS 368-371 respectively) design for Vλ CDR3 libraries.
- FIG. 39: Functional map of the pBS13 expression vector series.

- FIG. 40: Expression of all 49 HuCAL scFvs obtained by combining each of the 7 VH genes with each of the 7 VL genes (pBS13, 30°C.): Values are given for the percentage of soluble vs. insoluble material, the total and the soluble amount compared to the combination H3κ2, which was set to 100%. In addition, the corresponding values for the McPC603 scFv are given.
- Table 1: Summary of human immunoglobulin germline sequences used for computing the germline membership of rearranged sequences. (A) kappa sequences, (B) lambda sequences and (C), heavy chain sequences. (1) The germline name used in the various calculations, (2) the references number for the corresponding sequence (see appendix for sequence related citations), (3) the family where each sequence belongs to and (4), the various names found in literature for germline genes with identical amino acid sequences.
- Table 2: Rearranged human sequences used for the calculation of consensus sequences. (A) kappa sequences, (B) lambda sequences and (C), heavy chain sequences. The table summarized the name of the sequence (1),

the length of the sequence in amino acids (2), the germline family (3) as well as the computed germline counterpart (4). The number of amino acid exchanges between the rearranged sequence and the germline sequence is tabulated in (5), and the percentage of different amino acids is given in (6). Column (7) gives the references number for the corresponding sequence (see appendix for sequence related citations).

- Table 3: Assignment of rearranged V sequences to their germline counterparts.

 (A) kappa sequences, (B) lambda sequences and (C), heavy chain sequences. The germline genes are tabulated according to their family (1), and the number of rearranged genes found for every germline gene is given in (2).
- Table 4: Computation of the consensus sequence of the rearranged V kappa sequences. (A) (SEQ ID NO: 14), V kappa subgroup 1, (B) (SEQ ID NO: 15), V kappa subgroup 2, (C) (SEQ ID NO: 16), V kappa subgroup 3 and (D) (SEQ ID NO: 17), V kappa subgroup 4. The number of each amino acid found at each position is tabulated together with the statistical analysis of the data. (1) Amino acids are given with their standard one-letter abbreviations (and B means D or N, Z means E or Q and X means any amino acid). The statistical analysis summarizes the number of sequences found at each position (2), the number of occurrences of the most common amino acid (3), the amino acid residue which is most common at this position (4), the relative frequency of the occurrence of the most common amino acid (5) and the number of different amino acids found at each position (6).
- Table 5: Computation of the consensus sequence of the rearranged V lambda sequences. (A) (SEQ ID NO: 18), V lambda subgroup 1, (B) (SEQ ID NO: 19), V lambda subgroup 2, and (C) (SEQ ID NO: 20), V lambda subgroup 3. The number of each amino acid found at each position is

tabulated together with the statistical analysis of the data. Abbreviations are the same as in Table 4.

Table 6: Computation of the consensus sequence of the rearranged V heavy chain sequences. (A) (SEQ ID NO: 21), V heavy chain subgroup 1A, (B) (SEQ ID NO: 22), V heavy chain subgroup 1B, (C) (SEQ ID NO: 23), V heavy chain subgroup 2, (D) (SEQ ID NO: 24), V heavy chain subgroup 3, (E) (SEQ ID NO: 25), V heavy chain subgroup 4, (F) (SEQ ID NO: 26), V heavy chain subgroup 5, and (G) (SEQ ID NO: 27), V heavy chain subgroup 6. The number of each amino acid found at each position is tabulated together with the statistical analysis of the data. Abbreviations are the same as in Table 4.

In the case of the CDR3s, any sequence could be chosen since these CDRs were planned to be the first to be replaced by oligonucleotide libraries. In order to study the expression and folding behavior of the consensus sequences in E. coli, it would be useful to have all sequences with the same CDR3, since the influence of the CDR3s on the folding behavior would then be identical in all cases. The dummy sequences QQHYTTPP (see, for instance, positions 89-96 of SEQ ID NO: 28 and positions 88-95 of SEQ ID NO: 34) and ARWGGDGFYAMDY (positions 97-109 of SEQ ID NOS 35 & 36) were selected for the VL chains (kappa and lambda) and for the VH chains, respectively. These sequences are known to be compatible with antibody folding in E. coli (Carter et al., 1992).

1.5 Gene Design

The final outcome of the process described above was a collection of 14 HuCAL amino acid sequences, which represent the frequently used structural antibody repertoire of the human immune system (see FIG. 2). These sequences were back-translated into DNA sequences. In a first step, the back-translation was done using only codons which are known to be frequently used in E. coli. These gene sequences were then used for creating a database of all possible restriction endonuclease sites, which could be introduced without changing the corresponding amino acid sequences. Using this database, cleavage sites were selected which were located at the flanking regions of all sub-elements of the genes (CDRs and framework regions) and which could be introduced in all HuCAL VH, $V\kappa$ or $V\lambda$ genes simultaneously at the same position. In a few cases it was not possible to find cleavage sites for all genes of a subgroup. When this happened, the amino acid sequence was changed, if this was possible according to the available sequence and structural information. This exchange was then analyzed again as described above. In total, the following 6 amino acid residues were exchanged during this design (given is the name of the gene, the position

according to Kabat's numbering scheme, the amino acid found at this position as the most abundant one and the amino acid which was used instead):

VH2: T₃Q

VH6: S₄₂G

 $V\kappa 3$: E_1D , $I_{58}V$

 $V\kappa 4$: $K_{24}R$

Vλ3: T₂₂S

was carried out using the small hapten fluorescein bound to BSA (FITC-BSA) as antigen.

2.1 Cloning of the HuCAL VH3-Vk2 scFv Fragment

2.2 Construction of a Monovalent Phage-display Phagemid Vector pIG10.3

Phagemid pIG10.3 (FIG. 9) was constructed in order to create a phage-display system (Winter et al., 1994) for the H3κ2 scFv gene. Briefly, the EcoRI/HindIII restriction fragment in the phagemid vector pIG10 (Ge et al., 1995) was replaced

by the c-myc followed by an amber codon (which encodes an glutamate in the amber-suppresser strain XL1 Blue and a stop codon in the non-suppresser strain JM83) and a truncated version of the gene III (fusion junction at codon 249, see Lowman et al., 1991) through PCR mutagenesis.

2.3 Construction of H-CDR3 Libraries

Heavy chain CDR3 libraries of two lengths (10 and 15 amino acids) were constructed using trinucleotide codon containing oligonucleotides (Virnekas et al., 1994) as templates and the oligonucleotides complementing the flanking regions as primers. To concentrate only on the CDR3 structures that appear most often in functional antibodies, we kept the salt-bridge of R_{H94} and D_{H101} in the CDR3 loop. For the 15-mer library, both phenylalanine and methionine were introduced at position 100 since these two residues were found to occur quite often in human CDR3s of this length (not shown). For the same reason, valine and tyrosine were introduced at position 102. All other randomized positions contained codons for all amino acids except cystein, which was not used in the trinucleotide mixture.

The CDR3 libraries of lengths 10 and 15 were generated from the PCR fragments using oligonucleotide templates (SEQ ID NOS 4 & 5, respectively) O3HCDR103T (5'-GATACGCCGTGTATTATTGCGCGCGT (TRI)₆ GATTATTGGGGCCAAGGCACCCTG-3') (5'and O3HCDR153T GATACGGCCGTGTATTATTGCGCGCGT(TRI)₁₀ (TTT/ATG)GAT(GTT/TAT)TGGGGCCAAGGCACCCTG-3'), and primers (SEQ ID NOS 6 7, respectively) O3HCDR35 (5'-(5'-GATACGGCCGTGTATTATTGC-3') and O3HCDR33 CAGGGTGCCTTGGCCCC-3'), TRI where are trinucleotide mixtures representing all amino acids without cystein, (TTT/ATG) and (GTT/TAT) are trinucleotide mixtures encoding the amino acids phenylalanine/methionine and valine/tyrosine, respectively. The potential diversity of these libraries was 4.7 x 10^7 and 3.4 $\times 10^{10}$ for 10-mer and 15-mer library, respectively. The library cassettes were first synthesized from PCR amplification of the oligo templates in the presence of both primers: 25 pmol of the oligo template O3HCDR103T or O3HCDR153T, 50 pmol each of the primers O3HCDR35 and O3HCDR33, 20 nmol of dNTP, 10x buffer and 2.5 units of Pfu DNA polymerase (Stratagene) in

a total volume of 100 ml for 30 cycles (1 minute at 92°C., 1 minute at 62°C. and 1 minute at 72°C.). A hot-start procedure was used. The resulting mixtures were phenol-extracted, ethanol-precipitated and digested overnight with Eagl and Styl. The vector pIG10.3-scH3κ2cat, where the Eagl-Styl fragment in the vector pIG10.3-scH3κ2 encoding the H-CDR3 was replaced by the chloramphenicol acetyltransferase gene (cat) flanked with these two sites, was similarly digested. The digested vector (35 μg) was gel-purified and ligated with 100 μg of the library cassette overnight at 16°C. The ligation mixtures were isopropanol precipitated, air-dried and the pellets were redissolved in 100 ml of ddH2O. The ligation was mixed with 1 ml of freshly prepared electrocompetent XL1 Blue on ice. 20 rounds of electroporation were performed and the transformants were diluted in SOC medium, shaken at 37°C. for 30 minutes and plated out on large LB plates (Amp/Tet/Glucose)

4.4 Cloning of Low-copy Number Plasmid Vectors pCALO

A series of low-copy number plasmid vectors was constructed in a similar way using the p15A module M12 instead of the ColE1 module M14-Ext2. FIG. 35a is showing the functional maps and sequences of the vectors pCALO1 to pCALO3.

Example 5: Construction of a HuCAL scFv Library

5.1. Cloning of All 49 HuCAL scFv Fragments

All 49 combinations of the 7 HuCAL-VH and 7 HuCAL-VL consensus genes were assembled as described for the HuCAL VH3-Vκ2 scFv in Example 2 and inserted into the vector pBS12, a modified version of the pLisc series of antibody expression vectors (Skerra et al., 1991).

5.2 Construction of a CDR Cloning Cassette

For replacement of CDRs, a universal β-lactamase cloning cassette was constructed having a multi-cloning site at the 5'-end as well as at the 3'-end. The 5'-multi-cloning site comprises all restriction sites adjacent to the 5'-end of the HuCAL VH and VL CDRs, the 3'-multi-cloning site comprises all restriction sites adjacent to the 3' end of the HuCAL VH and VL CDRs. Both 5'- and 3'-multi-cloning site were prepared as cassettes via add-on PCR using synthetic oligonucleotides as 5'- and 3'-primers using wild type β-lactamase gene as template. FIG. 36 shows the functional map and the sequence of the cassette bla-MCS.

5.3. Preparation of VL-CDR3 Library Cassettes

The VL-CDR3 libraries comprising 7 random positions were generated from the PCR fragments using oligonucleotide templates $V\kappa1\&V\kappa3$, $V\kappa2$ and $V\kappa4$ and primers O_K3L_5 and O_K3L_3 (FIG. 37) for the $V\kappa$ genes, and $V\lambda$ and

primers (SEQ ID NO: 8) O_L3L_5 (5'-GCAGAAGGCGAACGTCC-3') and O_L3LA_3 (FIG. 38) for the V λ genes. Construction of the cassettes was performed as described in Example 2.3.

containing 90 µg chloramphenicol and 60 mM glucose) was inoculated overnight at 37°C. Next day the overnight culture was used to inoculate 30 ml LB medium containing chloramphenicol (30 μg/ml). The starting OD_{600nm} was adjusted to 0.2 and a growth temperature of 30.degree. C. was used. The physiology of the cells was monitored by measuring every 30 minutes for 8 to 9 hours the optical density at 600 nm. After the culture reached an OD_{600nm} of 0.5, antibody expression was induced by adding IPTG to a final concentration of 1 mM. A 5 ml aliquot of the culture was removed after 2 h of induction in order to analyze the antibody expression. The cells were lysed and the soluble and insoluble fractions of the crude extract were separated as described in Knappik & Pluckthun, 1995. The fractions were assayed by reducing SDS-PAGE with the samples normalized to identical optical densities. After blotting and immunostaining using the α-FLAG antibody M1 as the first antibody (see Ge et al., 1994) and an Fc-specific anti-mouse antiserum conjugated to alkaline phosphatase as the second antibody, the lanes were scanned and the intensities of the bands of the expected size (appr. 30 kDa) were quantified densitometrically and tabulated relative to the control antibody (see FIG. 40).

Example 7 Optimization of Fluorescein Binders

7.1. Construction of L-CDR3 and H-CDR2 Library Cassettes

A L-CDR3 library cassette was prepared from the oligonucleotide (SEQ ID NO: 9) template CDR3L (5'-TGGAAGCTGAAGACGTGGGCGTGTATTATT GCCAGCAG(TR5)(TRI)₄CCG(TRI)TTTGGCCAGGGTACGAAAGTT-3') and primer (SEQ ID NO: 10) 5'-AATTTCGTACCCTGGCC-3' for synthesis of the complementary strand, where (TRI) was a trinucleotide mixture representing all amino acids except Cys, (TR5) comprised a trinucleotide mixture representing the 5 codons for Ala, Arg, His, Ser, and Tyr.

A H-CDR2 library cassette was prepared from the oligonucleotide template CDRsH (SEQ ID NOS 11 & 12, respectively) (5'-AGGGTCTCG

AGTGGGTGAGC(TRI)ATT(TRI)₂₋₃(6)₂(TRI)ACC(TRI)TATGCG GATAGCGTGAAAGGCCGTTTTACCATTTCACGTGATAATTCGAAAAA CACCA-3'), and primer (SEQ ID NO: 13) 5'-TGGTGTTTTTCGAATTATCA-3' for synthesis of the complementary strand, where (TRI) was a trinucleotide mixture representing all amino acids except Cys, (6) comprised the incorporation of (A/G) (A/C/G) T, resulting in the formation of 6 codons for Ala, Asn, Asp, Gly, Ser, and Thr, and the length distribution being obtained by performing one substoichiometric coupling of the (TRI) mixture during synthesis, omitting the capping step normally used in DNA synthesis.